Comparison of Methods for Predicting the Compositional Dependence of the Density and Refractive Index of Organic-Aqueous Aerosols

Chen Cai[1,2], Rachael E. H. Miles[2,*], Michael I. Cotterell[2], Aleksandra Marsh[2], Grazia Rovelli[2,3], Andrew M. J. Rickards[2], Yun-hong Zhang[1] and Jonathan P. Reid[2]

[1] *The Institute of Chemical Physics, Key Laboratory of Cluster Science, Beijing Institute of Technology, Beijing 100081, People's Republic of China*

[2] *School of Chemistry, University of Bristol, Bristol, BS8 1TS, UK*

[3] *Department of Earth and Environmental Sciences, University of Milano-Bicocca, 20124 Milan, Italy*

** Corresponding author: Rachael.Miles@bristol.ac.uk, Tel: +44 117 331 7388*

Abstract

Representing the physicochemical properties of aerosol particles of complex composition is of crucial importance for understanding and predicting aerosol thermodynamic, kinetic and optical properties and processes, and for interpreting and comparing analysis methods. Here, we consider the representations of the density and refractive index of aqueous-organic aerosol with a particular focus on the dependence of these properties on relative humidity and water content, including an examination of the properties of solution aerosol droplets existing at supersaturated solute concentrations. Using bulk phase measurements of density and refractive index for typical organic aerosol components, we provide robust approaches for the estimation of these properties for aerosol at any intermediate composition between pure water and pure solute. Approximately 70 compounds are considered, including mono-, di- and tri-carboxylic acids, alcohols, diols, nitriles, sulphoxides, amides, ethers, sugars, amino acids, aminium sulphates, and polyols. We conclude that the molar refraction mixing rule should be used to predict the refractive index of the solution using a density treatment that assumes ideal mixing or, preferably, a polynomial dependence on the square root of the mass fraction of solute, depending on the solubility limit of the organic component. Although the uncertainties in

the density and refractive index predictions depend on the range of sub-saturated compositional data available for each compound, typical errors for estimating the solution density and refractive index are less than ±0.1 % and ±0.05 %, respectively. Owing to the direct connection between molar refraction and the molecular polarizability, along with the availability of group contribution models for predicting molecular polarizability for organic species, our rigorous testing of the molar refraction mixing rule provides a route to predicting refractive indices for aqueous solutions containing organic molecules of arbitrary structure.

1. Introduction

The composition of aerosol particles is most often complex in many applications or settings, consisting of mixtures of multiple compounds and varying in response to changes in the gas phase composition. For example, atmospheric aerosols consist of complex mixtures of many inorganic and organic species, responding to changes in ambient relative humidity (RH) and temperature that lead to changes in water partitioning between the condensed and gas phases and, thus, mass fractions of solutes within the particle. Changes in RH lead to changes in particle phase, size, density and refractive index (RI). These changes in droplet microphysical properties influence aerosol optical cross-sections (impacting climate radiative forcing), mass concentrations (affecting environmental air quality), and chemical reactivity (influencing atmospheric composition). In many instances, aerosol droplets exist in metastable supersaturated solution states or phases not accessible in the bulk-phase, precluding simple predictions of microphysical properties based on bulk phase studies and necessitating direct aerosol measurements. However, the microphysical properties of metastable states can be predicted using mixing rules providing that the densities and refractive indices of the pure solutes are known. Furthermore, these solute densities and refractive indices must correspond to the solute sub-cooled melt state, rather than the most stable crystalline form, for the application of mixing rules to be valid.[1] Accurate predictions of the aerosol core physicochemical properties, in particular density and RI, are essential to reconcile measurements made using different analytical techniques, for predictions of aerosol direct radiative forcing and for quantifying the activation of aerosol particles to form cloud droplets.[2] For example, particle sizes are reported as aerodynamic, electric mobility or optical, depending on the measurement technique and reconciling the different size distributions requires knowledge of aerosol particle density and RI.[3]

In this publication, we assess the accuracies of different mixing rules for predicting the densities and refractive indices of binary aqueous solutions containing a single organic component. Single particle and bulk phase measurements can be used to robustly test and refine mixing rules for predicting the densities and refractive indices of solutions of complex composition.[4] Although the accuracy of mixing rules has been considered in some detail for inorganic-aqueous solutions,[1,5] there has been significantly less consideration of the accuracy and validity of mixing rules for organic-aqueous solutions[6] with comparisons made for only a small selection of compounds (see, for example, references 7-9). Given the considerable chemical complexity in the organic fraction of ambient aerosol, it remains a significant challenge to interpret and represent the densities and refractive indices of their aqueous solutions in a consistent and accurate manner. Reported values of the RI for secondary organic aerosol (SOA) span from 1.36 to 1.66 with anthropogenic SOA considered more strongly scattering (and also absorbing) than biogenic SOA.[10] However, values of RI for SOA are mostly reported under only dry conditions, neglecting that hydration changes particle size and RI.[11,12] Furthermore, uncertainties in the measurement of scattering cross sections from nephelometers can be as large as ±10 %, and ±20 % from cavity ring down spectroscopy,[13,14] and ensuing uncertainties in the retrieved values of aerosol refractive index are typically of order ±0.05 (±3%).[15,16] Such a level of uncertainty often precludes validation of mixing rules for predicting mixed component RIs even though such accurate predictions are required. For example, Zarzana et al. concluded that an uncertainty of ±0.02 in RI for a 150 nm ammonium sulphate particle leads to an uncertainty of ~±7 % in radiative forcing; indeed, uncertainties in the imaginary part of the RI, which can also be represented by mixing rules, lead to even larger errors, often of tens of percent.[17]

To predict the densities of aqueous solutions of an organic component, it is common to assume a simple mass or volume additivity ideal mixing rule.[5] The latter requires knowledge of, or at least an estimate of, of the pure organic component density in the liquid state. Similarly, all mixing rules for RI require knowledge of, or an estimate of, of the pure organic component RI. A limited number of predictive tools have been considered for both pure component densities[2] and RIs,[12,18,19] frequently relying on group-contribution approaches. Although group contribution models are often based on a limited set of chemical functionalities with particular relevance to the chemical industry, density predictions have been compared with measurements for a limited number of organic components relevant to atmospheric aerosol. Moreover, the predictive tools for RI have largely been

developed for specific chemical systems and oxidation pathways[12] and their general applicability has been found to be limited.[20]

The interpretation of single aerosol droplet measurements in optical or electrodynamic traps often relies on knowledge of accurate relationships between RI, density and composition, relationships particularly important for interpreting studies of hygroscopic growth and oxidation chemistry.[21] Indeed, Reid and co-workers have shown that the composition dependence of the RI of a solution droplet and the pure solute melt density can be inferred with accuracies better than <0.05 % and ±2.5 %, respectively, for droplets where these relations were not known prior to measurement.[4,8,22-24] As such, these measurements could provide invaluable data for validating and refining models of density and RI for complex multicomponent aerosol droplets as a function of RH, particularly at RHs where the solutes exist in supersaturated states. From a comprehensive comparison of bulk phase measurements of density and RI under sub-saturated conditions for a wide range of organic compounds of varying functionality, we consider here the accuracy of a variety of methods for predicting these properties for aerosol in the supersaturated state. Such a comparison is of crucial importance for inferring droplet composition from single particle measurements of RI and particle size and, more generally, for predicting physicochemical properties for organic aerosol in any application or setting. Barley et al. and Liu et al. have provided comprehensive accounts of predictions of pure component densities for organic aerosol components and the refractive indices of dry mixtures of components, respectively.[2,6] We extend this to a consideration of the RH and water content dependence of these properties for mixed organic-aqueous aerosol.

In Section 2, we introduce the frequently used mixing rules for predicting the RIs of mixed component aerosol and assess the accuracy of these treatments based on binary systems for which RI and density values are available across the full composition range. We also consider the impact on these RI predictions of using an ideal mixing rule to calculate solution densities rather than measured values. In Section 3 we more fully explore the available methods for predicting both pure component organic densities, as well as accurately representing the densities of their supersaturated aqueous solutions. In particular, we consider the case of compounds that are only sparingly soluble in water. Extrapolations of pure melt densities from fitting the different density methods to bulk phase data are then compared with predictions using the method of Barley et al.[2] In Section 4, we use the most accurate density treatments determined in Section 3 in conjunction with a refractive index mixing rule to predict the RI of both the pure component organic melt and supersaturated organic-aqueous

aerosol droplets, comparing the results with both literature data and new measurements from single particle measurements using aerosol optical tweezers.

2. Refractive index mixing rules

Several RI mixing rule treatments for calculating the refractive indices of multicomponent mixtures can be found in the aerosol literature, with the three most common being the linear volume and linear mass weighting of RI, and the mole fraction weighting of molar refraction.

In the linear volume mixing rule, the effective RI, n_e, is calculated as a linear sum of the product $f_i n_i$ over all i components in a mixture:

$$n_e = \sum_i f_i n_i \qquad (1)$$

in which n_i is the RI of component i and f_i is its corresponding volume fraction. Similarly, the linear mass mixing rule describes n_e as a sum of the product $\phi_i n_i$, with ϕ_i the mass fraction of component i in the mixture:

$$n_e = \sum_i \phi_i n_i \qquad (2)$$

These two mixing rules have no underlying physical basis, yet have commonly been used for predicting aerosol refractive indices. By contrast, Liu and Daum[6] showed that the molar refraction mixing rule based on the mole fraction weighting of molar refraction is self-consistent with the Lorentz-Lorenz equation. Importantly, the Lorentz-Lorenz equation has a physical basis, relating the macroscopic optical (RI) or electrical (dielectric constant) properties of a pure slab of material to its microscopic molecular properties (molecular polarizability). The Lorentz-Lorenz equation describes the relationship between the material RI (n), the number density of molecules in the material (ρ_N) and the mean molecular polarizability of the substance (α):

$$\frac{n^2 - 1}{n^2 + 2} = \frac{\alpha \rho_N}{3} \qquad (3)$$

Using the relation for mass density, $\rho_M = M\rho_N / N_A$, equation (3) becomes:

$$\frac{n^2 - 1}{n^2 + 2} = \frac{N_A \alpha \rho_M}{3M} \qquad (4)$$

where M is the molecular weight of the substance and N_A the Avogadro number. This is often re-written in the form of equation (5), in which the quantity R is known as the molar refraction of the substance.

$$R = \left(\frac{n^2 - 1}{n^2 + 2}\right)\frac{M}{\rho_M} = \frac{N_A \alpha}{3} \qquad (5)$$

The Lorentz-Lorenz equation was originally derived for pure substances, whereas atmospheric aerosols are typically internal mixtures containing a variety of different chemical species. Therefore, the molar refraction treatment must be modified in order to be applicable to multi-component aerosol systems. An established way of treating inhomogeneous mixtures is by using an effective medium approximation, where the inhomogeneous mixture is modelled as a homogeneous material which has the effective (i.e. average) properties of all the species it contains, calculated through the use of mixing rules. In order to apply the molar refraction mixing rule to multi-component systems, effective values (denoted by a subscript e) of R, M, ρ_M and α must be calculated. As mass density, RI, molecular weight and molecular polarizability are related through the Lorentz-Lorenz equation, these properties must be considered together in an effective medium approximation, and the mixing rules chosen should be self-consistent with each other through the Lorentz-Lorenz equation.

A mixture which satisfies the Lorentz-Lorenz equation is assumed to be an ideal mixture in which the masses, volumes, moles and molecular polarizabilities of the different molecular constituents within the mixture are conserved following the mixing process. This physical picture forms the basis of the self-consistent mixing rules used to calculate R_e, M_e, ρ_{Me} and α_e.[6] Polarizability additivity requires that the polarization $P=N\alpha_e$ of a mixture be equal to the sum of the polarizabilities of each component in the mixture and thus the effective polarizability of a mixture is given by:

$$\alpha_e = \sum_i \frac{N_i}{N}\alpha_i = \sum_i \frac{y_i N_A}{y N_A}\alpha_i = \sum_i x_i \alpha_i \qquad (6)$$

in which y_i, x_i and N_i are the number of moles, mole fraction and number of molecules of component i in the mixture, respectively. Substitution of equation (6) into (5) yields the mole fraction mixing rule for the molar refraction.

$$R_e = \frac{N_A \alpha_e}{3} = \frac{N_A}{3}\sum_i x_i \alpha_i = \sum_i x_i R_i \qquad (7)$$

From the calculation of the effective molar refraction, R_e, the effective RI (n_e) can be determined through equation (5) provided that the effective molecular weight (M_e) and effective mass density (ρ_{Me}) can be calculated. From conservation of mass, M_e is calculated by:

$$M_e = \frac{m}{y} = \sum_i \frac{y_i M_i}{y} = \sum_i x_i M_i \qquad (8)$$

in which m is the total mass of the mixture. The self-consistent mixing rule for calculating the effective mass density is based on the principles of mass and volume additivity and is given by:

$$\frac{1}{\rho_{Me}} = \frac{1}{M_e} \sum_i \frac{M_i x_i}{\rho_{mi}} = \sum_i \frac{\phi_i}{\rho_{mi}} \qquad (9)$$

where ϕ_i is the mass fraction of species i in the mixture and ρ_{mi} is the pure component density. It is important to note the pure component densities and refractive indices used in RI mixing rule calculations to predict multi-component mixture properties should be those of the chemical species in its liquid or melt state at room temperature, not the solid/crystalline state.[1]

2.1. Assessment of refractive index mixing rules for binary aqueous-organic systems

To test the accuracy of each of the aforementioned three mixing rules at predicting the RI of multi-component mixtures, literature and experimental RI data for binary aqueous solutions containing different mass fractions of miscible organic liquids were compared with mixing rule model predictions for 21 different compounds, listed in Table S1.[25-33] The organic species considered included alcohols, nitriles, carboxylic acids, diols, sulphoxides, amides, ethers and polyols. As the molar refraction mixing rule requires knowledge of mixture density in order to predict RI, data sets were selected based on the availability of paired measurements of RI and density for each different mass fraction of the water/organic mixture. All studies also included measurements of the pure component densities, which are necessary for both the molar refraction mixing rule and the linear volume mixing rule if composition in mass or mole fraction needs to be converted to volume fraction. The choice of literature data was limited only to experimental studies performed on mixtures of organic species with water due to the importance of water as a solvent in atmospheric aerosols. All refractive indices were reported at a wavelength of 589 nm and at a temperature of 25 °C, the same temperature as for

the density measurements. In total, data for 290 different binary aqueous-organic mixture compositions were studied, with variation in both the identity of the organic species and the mass fraction of solute.

Mixture refractive indices were calculated using the linear mass and linear volume mixing rules using equations (1) and (2) respectively, with equation (10) used in the latter mixing rule to convert solute mass fraction to volume fraction:

$$f_s = \frac{\phi_s \rho_w}{\phi_s(\rho_w - \rho_s) + \rho_s} \qquad (10)$$

To be self-consistent, the pure component densities for water (subscript *w*) and solute (subscript *s*) were assigned the values reported in each literature study. For the molar refraction mixing rule, equations (5) and (7) were used to calculate the effective molar refraction and thus the mixture RI, calculating M_e using equation (8) and taking ρ_{Me}, ρ_w, n_w, ρ_s and n_s as reported in the relevant literature study. Figure 1 shows both the percentage and absolute difference between the predicted refractive indices with variation in mass fraction of solute in the aqueous mixture using each of the three RI mixing rules, and the experimentally measured values for each organic species. It is clear from both figures that the molar refraction mixing rule consistently outperforms the linear mass and linear volume mixing rules for all compounds studied, with the molar refraction mixing rule able to predict RI with an accuracy of better than ± 0.25 % or ± 0.002. This is better than the level of accuracy achieved in the measurement of refractive indices by typical experimental techniques that measure the optical properties of aerosol ensembles although remains less accurate than measurements made using single particle techniques.[14,15] It is also worth noting that the difference between the measured refractive indices and the molar refraction predictions for the 21 different compounds are evenly distributed around zero, with the model sometimes over-predicting the RI and sometimes under-predicting it. With the exception of one system (glycerol-water), the linear mass and linear volume mixing rules consistently over-predict the solution RI.

2.2. Accuracy of the molar refraction mixing rule when the mixture densities are calculated using a mixing rule

As shown above, the molar refraction mixing rule combined with accurate solution density measurements is able to predict the RI of a mixture to a good level of accuracy. However, such measurements of mixture density are not always readily available, especially with aerosols often existing in a supersaturated solute state inaccessible to bulk measurement techniques. In these instances, some form of density mixing rule is required to estimate the solution density at a given mass fraction composition. The self-consistent density treatment suggested by Liu and Daum[6] for use with the molar refraction mixing rule is based on an assumption of ideal mixing, an assumption that does not represent the behaviour of many aqueous-solute systems accurately. Indeed, the aqueous-organic liquid mixtures considered in the previous section show significant non-ideal behaviour upon mixing, as demonstrated below.

To determine the accuracy with which the self-consistent (ideal mixing) rule can predict the densities of aqueous-organic mixtures, equation (9) was used to calculate the density for each solute mass fraction value for each of the 21 compounds considered in Figure 1. These values were then compared with the experimentally measured mixture densities reported in the literature for compositions intermediate between the two pure component extremes. Figures 2(a) and 2(b) show the results as the percentage difference and absolute difference, respectively.

For aqueous mixtures of all 21 organic compounds, Figures 2(a) and 2(b) show that the ideal mixing density treatment over predicts the experimental mixture density. Considering trends for individual organic compounds, the largest maximum observed percentage deviation between predicted and measured densities is 3.7% which occurs for aqueous mixtures of N,N-dimethylacetamide, with the smallest maximum deviation of 1.1% seen for aqueous glycerol mixtures. In absolute density terms, the largest maximum deviation in density is again observed for aqueous mixtures of N,N-dimethylacetamide (0.036 $g.cm^{-3}$), with the smallest maximum absolute deviation observed for 1,3-propanediol (0.012 $g.cm^{-3}$).

Figures 3(a) and 3(b) report the level of agreement between measured and predicted refractive indices estimated from the molar refraction mixing rule when ideal mixing densities are used (equation (9)) rather than the actual experimentally measured density values. In contrast with Figure 1, the molar refraction mixing rule now consistently over predicts the mixture RI, with the largest maximum deviation between prediction and measurement of 0.016 (1.15 %) occurring for N,N-dimethylacetamide, and the smallest maximum deviation of 0.005 (0.33 %) for 1,3-propanediol. This level of agreement in RI is similar to that observed for N,N-

dimethylacetamide and 1,3-propanediol when mixture refractive indices are predicted using either the linear mass or linear volume mixing rules. This clearly suggests that the increased accuracy of the molar refraction mixing rule, over the mass and volume weighted alternatives, comes from its inclusion of accurate mixture density measurements in the calculation of the RI, demonstrating the importance of a robust and reliable method for predicting mixture density if experimental data are not available.

2.3. Application of refractive index mixing rules to surrogates of atmospheric organic aerosol

So far, our limited consideration of 21 aqueous organic solute systems has suggested that the molar refraction mixing rule is by far the most accurate method to predict solution refractive indices compared with the linear mass and linear volume mixing rules when accurate mixture density measurements are available, along with pure solute melt densities and refractive indices. Even when using an ideal mixing rule to calculate mixture densities from pure component values, the molar refraction treatment achieves the same level of agreement as the linear mass and linear volume mixing rules, while benefiting from having a robust, physical basis. Figure 4 summarises the error associated with predictions of density and RI for all of the 21 aqueous-organic systems, assuming ideal mixing when estimating the mixture density and using the molar refraction mixing rule for predicting the mixture RI. There is a clear correlation between the uncertainties in the two quantities, allowing an upper limit to be placed on the accuracy with which the mixture density must be determined if a specified level of uncertainty in the RI prediction from molar refraction calculations is required.

For the aqueous-organic liquid mixtures considered so far, measurements of both the pure component and mixture densities and refractive indices are trivial using bulk techniques. However, a large fraction of both inorganic and organic solutes of relevance to atmospheric aerosols are crystalline solids at room temperature and form supersaturated salt solutions in the aerosol state that are inaccessible to bulk measurement methods. The required values of densities for the solute melts at room temperature are largely unavailable for such compounds as they cannot be directly measured; a similar problem is encountered for the solute melt refractive indices. This potentially compromises the application of any of the RI mixing rules considered above as all require a value for the solute melt RI, with the linear volume and molar refraction mixing rules also requiring a value for the solute melt density, and with the latter further requiring mixture densities. Thus, for the

application of RI mixing rules to typical atmospheric aerosol compositions, robust methods are required for not only estimating the solute melt refractive indices, but also solute melt and mixture densities.

Having clearly demonstrated the superior accuracy of the molar refraction mixing rule predictions of aqueous-organic solution refractive indices, we now consider how this model can be applied to systems of atmospheric interest for determining both supersaturated solution and solute melt refractive indices. We first investigate different methods for predicting the solute melt and mixture densities, benchmarking each prediction technique against a large data set of sub-saturated solution density measurements taken from both new experimental studies and existing literature data, with the aim of determining the method that provides the most reliable prediction of solution density (Section 3). As discussed above, this is essential for accurate application of the molar refraction mixing rule. We then use the molar refraction mixing rule with this optimal density treatment and experimental sub-saturated solution RI data to predict solute melt refractive indices, enabling the RI of supersaturated solute solutions at all compositions to be calculated with confidence (Sections 4 and 5).

3. Predicting pure component melt densities and aqueous-solute mixture densities in the supersaturated regime

In this section, we evaluate several methods for predicting densities of atmospherically relevant aqueous-organic mixtures, using agreement with sub-saturated mixture density data for a new set of 47 organic compounds as a means of benchmarking each approach. Table S3 lists the 47 compounds considered, which includes branched and straight chain dicarboxylic acids, sugars, amino acids, aminium sulphates, monocarboxylic acids, tricarboxylic acids and substituted dicarboxylic acids, all of which are crystalline solids at room temperature. The three different density fitting methods considered are: i) the self-consistent density ideal mixing rule; ii) an apparent molar volume treatment based on that previously used by Clegg and Wexler; and iii) simple polynomial fits of second, third and fourth order to sub-saturated solution density measurements. Each method is used to predict the solute melt density for each of the organic compounds, which can be compared with predictions from UManSysProp,[Error! Reference source not found.] a group contribution model that predicts sub-cooled liquid densities (i.e. melt values) at room temperature based on compound molecular structure.

Sub-saturated aqueous-organic solution density data for the 47 compounds used in the benchmarking process were taken either from the literature, or from new experimental measurements performed as part of this study (data provided in Table S4). In the latter case, a series of aqueous solutions containing each compound were made at mass fractions of solute up to the bulk solubility limit, and the solution density and RI (at wavelength 589 nm) measured using a vibrating capillary density meter (Mettler Toledo Densito, accuracy ±0.001 g.cm^{-3}) and a refractometer (Misco Palm Abbe, accuracy ±0.0001), respectively. We checked the calibration of both instruments using distilled water (Purite Select Fusion) prior to each series of measurements. In total, density and RI measurements of 584 different aqueous-organic solutions were considered.

3.1. Fitting of bulk data to self-consistent density ideal mixing rule

For the two component organic-aqueous mixtures studies here, the ideal mixing rule for density (equation (9)) can be re-written in full in terms of solute mass fraction as:

$$\frac{1}{\rho_{Me}(1-\phi_s)} = \frac{\phi_s}{\rho_s(1-\phi_s)} + \frac{1}{\rho_w} \qquad (11)$$

Subsaturated density data for aqueous mixtures of each of the 47 organic compounds were plotted in the form $1/(\rho_{Me}(1-\phi_s))$ against $\phi_s/(1-\phi_s)$ and a linear fit determined with the y-intercept constrained to the value $1/\rho_w$. The value of ρ_w was taken as the pure water density measured by the Densito density meter in each series of measurements, or given in the relevant literature study. From equation (11), the gradient of the linear fit is equal to $1/\rho_s$, i.e. the reciprocal of the pure organic solute melt density. The 95% confidence intervals of the fit were included on each trend line allowing the uncertainty in the melt density to be calculated. Figure 5(a) shows an example of a fit of equation (11) to data for DL-Malic acid.

Figure 5(b) shows the difference between the experimentally measured density values and those predicted by the ideal mixing treatment using the solute melt density calculated from the linear fit for aqueous mixtures of the dicarboxylic acids, sugars, amino acids, mono- and tricarboxylic acids, substituted dicarboxylic acids and aminium sulphates. The figure shows that there are distinct differences in the accuracy with which the ideal mixing treatment can represent the density trend for aqueous mixtures of different organics. Mixtures containing the sugars and the aminium sulphates show differences that indicate clear departures from ideal

behaviour even at sub-saturated concentrations, with the ideal mixing treatment failing to represent the trend in the experimental values. These classes of compounds have errors of up to 0.021 g.cm^{-3} when predicting the densities of the most concentrated sub-saturated solutions, corresponding to an absolute error in RI of 0.0083 using the relationship given in Figure 4. Better agreement is observed for the dicarboxylic acids and all of the amino acids apart from one (L-Lysine). The majority of these compounds have lower solubility limits compared to the sugars and aminium sulphates, and it may be that the ideal mixing rule is better able to represent these compounds as their lower solubility means bulk measurements are not able to access states where the solutions start to become appreciably non-ideal.

3.2. Predicting mixture densities using the apparent molar volume

A second method for parameterising aqueous-organic solution density is to represent it in terms of its apparent molar volume, V^{ϕ}, (cm^3 mol^{-1}), defined as:

$$V^{\phi} = \frac{M_s}{\rho_{Me}} - 18.0152\frac{(1-x_s)}{x_s}\frac{(\rho_{Me}-\rho_w)}{\rho_{Me}\rho_w} \qquad (12)$$

where the term 18.0152 comes from the molar mass of pure water. Clegg and Wexler have previously used an apparent molar volume treatment to parameterise the variation in density with mass fraction of solute for 10 aqueous-inorganic systems (H_2SO_4, HNO_3, HCl, Na_2SO_4, $NaNO_3$, NaCl, $(NH_4)_2SO_4$, NH_4NO_3 and NH_4Cl), fitting critically evaluated sub-saturated and supersaturated literature density data.[1] The authors reported that the variation in V^{ϕ} with solute mass fraction, ϕ_s, could be well-described by an equation combining a mole fraction weighted Debye-Hückel term with a polynomial fit in terms of mass fraction of solute, giving an equation of the form:

$$V^{\phi} = \left(V^{\phi\infty} + C_{DH}\right) + b'\phi^{0.5} + c'\phi^{0.75} + d'\phi + e'\phi^{1.5} + f'\phi^{2} + g'\phi^{2.5} + h'\phi^{3} + k'\phi^{3.5} \qquad (13)$$

in which $V^{\phi\infty}$ is the infinite dilution apparent molar volume, C_{DH} is the Debye-Hückel term, and *b', c', d'* etc. are fitting constants. This equation, previously used by Tang et al.,[5,34,36] has no underlying physical basis, but was chosen by the authors as the most reliable way of parameterising the data. The independent variable was chosen to be a function of $\phi^{0.5}$ rather than ϕ as this reduced the extent of the extrapolation required to determine

the melt density. This is the same method employed by Cai et al.,[8] who parameterised the density of aqueous-organic solutions by fitting a fourth order polynomial in the $\phi^{0.5}$ domain to sub-saturated solution measurements for dicarboxylic acids.

Several problems arise when trying to apply the apparent molar volume treatment to predict the densities of the aqueous-organic systems studied here. Firstly, Clegg and Wexler did not always include all the ϕ terms given in equation (13) to achieve the best-fit of the model to the literature measurements for each of the aqueous-inorganic systems; instead, they adopted the form that best represented the data. For example, the fit to the experimental aqueous NaCl data omits the $\phi^{0.75}$, $\phi^{2.5}$, ϕ^{3} and $\phi^{3.5}$ terms, including instead a $\phi^{1.75}$ term, while the expression for $(NH_4)_2SO_4$ omits the $\phi^{0.5}$, $\phi^{0.75}$ and $\phi^{2.5}$ terms. In order for a density mixing rule to be confidently implemented for the aqueous-organic mixtures, the generality of the functional form to robustly represent all systems must be assumed. This is particularly important, as there are considerable variations in the mass fraction of solute range over which data are available, a consequence of the wide differences that can exist in compound solubility. The second problem is that equation (13) requires knowledge of an infinite dilution apparent molar volume for all solutes and a Debye-Hückel term for ionic salts, both of which are not readily available in the literature for the wide range of compounds of atmospheric interest.

The work of Clegg and Wexler demonstrated that the variation in V^{ϕ} with $\phi^{0.5}$ was near-linear for the inorganic electrolyte solutions considered, suggesting that it may be possible to neglect the higher order terms in equation (13) and combine the $V^{\phi\infty}$ and C_{DH} terms in to a single fit constant as a first approximation for inorganic solutions. If this general form for $V^{\phi\infty}$ can also be applied to the aqueous-organic solutions studied here, it might be possible to use the truncated version of equation (13) to predict the supersaturated solution and solute melt densities. Figure 6 shows the variation in apparent molar volume with the square root of the solute mass fraction for sub-saturated aqueous solutions of three different organic solutes with similar mass fraction bulk solubility limits; malonic acid, urea and maltose. In contrast with the aqueous-inorganic systems considered by Clegg and Wexler, the variations in V^{ϕ} with $\phi^{0.5}$ for the aqueous-organic systems are observed to be highly variable, depending on the identity of the organic compound in question. Only malonic acid is observed to show a linear trend in V^{ϕ} with $\phi^{0.5}$, comparable with that seen for the aqueous-inorganic solutions.

The lack of consistency between the trends in apparent molar volume observed for the different organics precludes not only the use of a truncated version of equation (13), but would also discount a fit in terms of apparent molar volume altogether. The stark differences between the shapes observed for the different compounds show that there is no simple relation relating V^{ϕ} with solute mass fraction that could accurately represent all the organic species of interest. In addition, as the trends are so different there would be an added complication when looking at compounds with different solubility limits: there would be no way of assessing if the correlation observed for sub-saturated solution concentrations was representative of the trend extending into the supersaturated concentration range. For example, a completely different trend would be fitted to the data for urea if the bulk solubility limit corresponded to a square root of mass fraction of solute of 0.4 rather than 0.7. For these aforementioned reasons, the apparent molar volume treatment was discounted as a means of predicting supersaturated and pure component melt densities for aqueous-organic systems.

3.3. Density predictions based on a simple polynomial fit

The final method assessed was the fitting of sub-saturated solution densities for aqueous-organic mixtures using a simple polynomial fit, with second, third and fourth order variants tested. Similar to equation (13) for the apparent molar volume treatment, use of a polynomial fit has no physical basis but is a common method employed when an extrapolation of available data in to a new region must be made. As in the work of Cai et al.,[8] sub-saturated density data were plotted as a function of $\phi^{0.5}$ to reduce the range for extrapolation to the pure melt value. For comparison, each aqueous-organic mixture data set was fitted with a second order polynomial in $\phi^{0.5}$, a third order polynomial in $\phi^{0.5}$ and a fourth order polynomial in $\phi^{0.5}$, with the intercept constrained to the value of pure water as measured by the density meter or given in the relevant literature study. Error bounds corresponding to 95% confidence intervals on each fit were also calculated. Figures 7(a) to 7(c) show the difference (residual) between the densities predicted by each polynomial fit equation and the measured solution densities for the 47 aqueous-organic systems under consideration.

It is clear from Figure 7(a) that the second order polynomial fit is the least able to represent the observed variation in density with square root of mass fraction of solute, particularly for the case of the aqueous-sugar solutions studied. For these compounds, density differences as large as 0.008 g cm^{-3} were observed between

the predicted and measured densities, significantly larger than the uncertainties associated with typical solution density measurements. Figures 7(b) and 7(c) show that the density differences between the experimental measurements and the third and fourth order polynomial density fits were very similar and within the uncertainty limits associated with the measurements of solution density made using the density meter. However, the error envelopes (defined by the 95% confidence intervals) on the fourth order fits were consistently larger than those for the third order polynomial fits, giving much larger uncertainties in predicted values for melt densities and supersaturated solution densities (Figure 8(a)). For some solutes, the error envelopes when fitting with a fourth order polynomial included ρ_{melt} values that were less than the density of pure water (and even negative in some cases) for 14 of the 47 compounds tested. This occurred exclusively for compounds with the lowest bulk solubilities where an extrapolation had to be made to the melt density from a very small range of solute mass fraction data, highlighting the importance of the fitting range on the accuracy of the parameterisation. Based on the above evidence, the third order polynomial fit was selected in preference to the fourth order fit as the polynomial best able to represent the sub-saturated density data.

3.4. Comparison of treatments for the case of low solubility compounds

As discussed in section 3.3, the level of agreement between experimental measurements and third order polynomial density predictions, along with the associated extrapolated uncertainties in ρ_{melt}, depends on the mass fraction of solute range of the bulk sub-saturated data available for fitting. Figure 7(b) shows that the residual between the density measurements and the third order polynomial fit is generally less than 0.001 $g.cm^{-3}$, a limit equivalent to the uncertainty on the density meter measurements themselves and substantially less than the errors when applying the ideal mixing density treatment (Figure 5(b)). This finding is not unexpected given that ideal mixing does not accurately reflect the properties of complex or concentrated solute mixtures. At first glance, the third order polynomial fit treatment appears to provide the most accurate parameterisation of density based on sub-saturated bulk data, permitting supersaturated and melt densities to be predicted. However, Figure 8(a) indicates a problem with the third order polynomial treatment: large uncertainties in the melt density can arise for relatively insoluble compounds for which bulk solutions cannot be prepared up to a high mass fraction of solute. It is clear that the precision in the melt density for compounds with bulk solubility

limits of less than $\phi \sim 0.4$ is poorer for the third order polynomial fit than that seen for the ideal mixing treatment by up to an order of magnitude. It is important to note that this higher precision for the ideal mixing treatment does not mean that the ρ_{melt} density predictions for these compounds are necessarily more accurate than those for the third order polynomial treatment, rather that the uncertainties on the extrapolated melt values are much less. While the primary goal of a density fitting treatment should be achieving an accurate density description, we must consider the precision of a parameterisation if meaningful trends or relationships are to be extracted from experimental data; a parameterisation with a large uncertainty is unlikely to allow substantive conclusions to be drawn. It is therefore recommended that the ideal mixing treatment is used to predict the density for insoluble compounds where bulk density data are unavailable over a wide range of mass fraction of solute, with the third order polynomial fit used for more soluble compounds when a wider concentration range of bulk solutions can be accessed. We now consider the delineating concentration at which the change between the parameterisations should occur.

Figure 8(b) reports the dependence of the estimated value of ρ_{melt} on the mass fraction of solute range over which a fit is performed using a third order polynomial fit to sub-saturated density data. Eight compounds are considered (sucrose, lactic acid, MEAS, DEAS, sorbitol, glucose, maltose and citric acid), all with mass fraction bulk solubility limits of between ~0.59 and ~0.8. These are the most soluble compounds considered in this study allowing us to investigate the influence of the truncation in data on the precision of the estimated melt density. The ρ_{melt} value predicted by a third order polynomial fit to the sub-saturated data plotted as a function of $\phi^{0.5}$ was repeatedly calculated for each compound, systematically removing data points corresponding to higher values for solute mass fraction from the fit. Subsequently, each prediction of ρ_{melt} was compared with the value obtained when all available solute mass fraction data was fitted. Dashed lines showing a ± 2% deviation from the ρ_{melt} density from the fit to the full range of solute mass fraction data are also included. Figure 8(b) clearly shows that fitting a third order polynomial to bulk density data that extends above a ϕ of 0.4 gives ρ_{melt} values representative (i.e. to within 2%) of the value attained when using a much wider range of solute mass fraction data, with the exception of MEAS. Considering this conclusion and the observations from Figure 8(a), the precision of the ideal mixing density treatment is equivalent to that from the third order polynomial fit when the bulk solubility limit of a compound is greater than ~0.4 mass fraction of solute. We therefore recommend that the ideal mixing treatment should be used for predicting density for

organic compounds with bulk solubility limits below ϕ = 0.4, while the third order polynomial treatment should be used for organic compounds with solubilities of ϕ = 0.4 or higher.

3.5. Comparison of solute melt densities fitted from experimental data with predictions

Melt density estimates made using the ideal mixing treatment and the third order polynomial fit were compared with predictions from UManSysProp which uses a group contribution method to predict the sub-cooled liquid density of organic species based on their molecular structure.[Error! Reference source not found.] Compounds are described using a SMILES string, which is then broken down into individual functional groups based on a SMARTS library. Each functional group has an assigned 'contribution' to the sub-cooled liquid density and these are combined according to theoretical equations to predict the melt density for a given organic compound. Although not validating our approach, comparison of the extrapolated melt densities with these model predictions allows an assessment of the consistency of the two approaches. Barley et al. assessed seven different theoretical methods for combining the functional group contributions in order to predict liquid density and compared the results of each with experimental liquid density data for 56 multifunctional organic compounds (228 data points in total).[2] They found that the group contribution method of Schroeder used in conjunction with the Rackett equation with critical properties by Nannoolal gave the best agreement with the experimental data set. The mean absolute error between the melt density predictions and the experimental data when results from all seven estimation methods were considered was 5%; we take this as the estimated uncertainty in the UManSysProp liquid density predictions in this study.

UManSysProp[Error! Reference source not found.] was used to predict the sub-cooled liquid (melt) densities of the 47 different organic compounds used in this study at a temperature of 293.15 K using the Schroeder/Nannoolal treatment recommended by Barley et al.[2] It was not possible to predict the melt densities of the amino acids and the aminium sulphates due to the presence of the aliphatic nitrogen species, which the model does not take into account. Problems were also encountered when simulating the liquid densities of the sugars as they can exist in both straight chain and ring conformations. As UManSysProp predicts sub-cooled liquid density using a group contribution method based on molecular structure, the melt density prediction is different depending on whether the SMILES string corresponds to the straight chain or to the ring form of a compound. The

difference between the melt densities predicted for the chain/ring structure can be significant (eg. 1.41 g.cm^{-3} for glucose in a ring conformation vs. 1.26 g.cm^{-3} in open conformation), falling outside the accepted ± 5% error on the predictions. In this study, all melt densities were predicted using the ring conformation of the sugars where two forms are known to exist. Table S5 lists the SMILES strings used for each compound and the corresponding UManSysProp predictions.

Figure 9 shows the levels of agreement between the solute melt densities predicted by UManSysProp, the self-consistent ideal mixing density treatment and the third order polynomial fit, along with their associated uncertainties. Data are shown in terms of the maximum mass fraction of solute achievable in the bulk density measurements to illustrate the extent of data extrapolation required to calculate ρ_{melt}. For the dicarboxylic acids, excellent agreement is observed between the melt densities predicted by UManSysProp and the self-consistent ideal mixing density rule, with all ideal mixing points lying within the error bar of the group contribution predictions. Consistency is marginally poorer between the predicted densities and the third order polynomial fit values , which also have large uncertainties particularly when the maximum in the mass fraction of solute data is low; the level of consistency improves as the mass fraction of solute achievable in the bulk increases, as expected from the discussion in section 3.4. For the majority of the dicarboxylic acids considered, the ideal mixing density treatment is the most appropriate for predicting the density given their low solubility limits.

For the sugars (Figure 8(b)), the self-consistent ideal mixing density treatment consistently predicts higher melt densities than the third order polynomial treatment, and both are higher than the result from UManSysProp. For some compounds, the UManSysProp melt density prediction is less than the density measured for the highest mass fraction of solute solution achievable in the bulk, suggesting a problem with the group contribution method for these compounds. Figure 8(c) considers the results for the mono- and tricarboxylic acids, as well as malic and tartaric acids. Again, there is generally good agreement between the UManSysProp predictions and the melt density calculated by the self-consistent ideal mixing density treatment, while the agreement with the third order polynomial ρ_{melt} values improves as the bulk solubility of compounds increases. As with the dicarboxylic acids, UManSysProp predictions agree with the extrapolated melt densities to within the stated uncertainty of the group contribution model.

4. Determining the melt refractive index using the molar refraction mixing rule

Once the appropriate density treatment has been identified for a given organic compound based on the discussion in section 3.4, this density treatment can be used in combination with the molar refraction mixing rule to fit experimental sub-saturated RI data and predict the solute melt RI. Using this value, the RI of supersaturated aqueous-organic solutions can then be predicted at any concentration.

For a binary mixture of water and solute, the effective molar refraction, R_e, is given by:

$$R_e = (1 - x_s)R_w + x_s R_s \qquad (14)$$

in which R_w is the molar refraction of water and R_s is the solute molar refraction. The value of R_w was calculated using equation (5) taking n_w as 1.333 at λ = 589 nm and ρ_w as the value measured by the density meter or given in the literature study for pure water for each of the aqueous-organic mixtures studied. In the fitting process, trial solutions of the solute melt RI, n_s, were used in equation (5) to calculate R_s, with the solute melt density determined from the appropriate density parameterisation.[2] Equation (14) was then used to calculate the corresponding R_e for the full range of solute mass fractions from ϕ = 0 to ϕ = 1, with R_e then converted to n_e using equation (5) in conjunction with the predicted mixture densities (Table S6) and the effective molecular masses (equation (8)). The calculated values of n_e were then compared with the measured refractive indices for the sub-saturated solutions and a least-squares fit residual calculated. The value of n_s was varied until the best agreement (minimum residual-squared) was found between the experimental data and the molar refraction mixing rule predictions. The Excel GRG non-linear engine, a solver built in to Microsoft Excel that uses the Generalized Reduced Gradient Algorithm for optimising non-linear problems, was used to perform the least-squares fitting routine. Error bounds on the value of the melt RI were calculated based on the uncertainty in the density mixing rule parameterisation. Table S6 gives the determined melt refractive indices for all 47 organic compounds studied and their associated uncertainties.

Figure 10(a) shows the difference between the measured refractive indices of sub-saturated aqueous solutions of all 47 compounds considered in this study, and the refractive indices predicted by the molar refraction mixing rule, plotted as a function of mass fraction of solute. Predicted RIs were calculated using the best-fit density treatment determined for each compound based on the criterion discussed in section 3.4. For all compounds, the largest deviation between the molar refraction mixing rule predictions and the sub-saturated

RI measurements is less than 0.001 across the entire MFS range. This compares extremely favourably with the 0.002 maximum deviation observed in Figure 1(b), when the molar refraction mixing rule was used to predict the RI of the 21 liquid organic-aqueous solutions where both the pure component densities and RIs as well as the mixture densities were known. Figure 10(a) demonstrates that the form of the molar refraction mixing rule can accurately represent the observed trend in solution RI with variation in mass fraction of solute.

Figure 10(b) shows the uncertainty in the melt RI estimated by fitting the molar refraction mixing rule to sub-saturated RI data, plotted as a function of the maximum solute mass fraction achievable in the bulk. A clear demarcation between compounds which fall in to the two density treatment regimes can be seen, with the ideal mixing density treatment (maximum $\phi < 0.4$) giving, in general, smaller uncertainties in melt RI than the third order polynomial density treatment (maximum $\phi \geq 0.4$). For all but three of the 47 compounds tested, the absolute uncertainty in the melt density is less than 0.007, significantly better than the 0.02 uncertainty typically reported by some experimental measurement techniques.[14,15]

5. Comparison of experimental refractive index measurements with RI predictions from the molar refraction mixing rule

The variation in solution RI with mass fraction of solute predicted by the molar refraction mixing rule was tested in the supersaturated solute regime for two compounds (citric acid and malonic acid) with measurements made using aerosol optical tweezers (AOT). Density parameterisations and pure component refractive indices for use in the molar refraction mixing rule were determined from bulk phase, subsaturated measurements in the manner already described. The AOT technique has been described in detail previously and will only be briefly reviewed here; the reader is directed to references 37-39 for a full account of the method and its capabilities. In an AOT measurement, a single particle is held using a focussed 532 nm laser beam inside a trapping chamber, the relative humidity of which is regulated through the mixing of wet and dry nitrogen gas flows (giving an accessible RH range of 0 – ~90 % RH). The trap RH is measured using a capacitance probe with an accuracy of ± 2% at < 80% RH. Excitation of molecular species within the droplet by the trapping laser leads to spontaneous and stimulated Raman scattering, with light inelastically backscattered by the droplet collected and dispersed using a spectrograph. Cavity Enhanced Raman Spectroscopy (CERS) is used to probe

the physical, chemical and optical properties of the aerosol. By comparing the wavelengths at which stimulated Raman peaks occur with Mie theory, the particle radius, RI and wavelength dispersion can be retrieved. By varying the RH within the trapping cell, supersaturated solute states can be accessed in the trapped droplet that are unattainable in bulk solutions, allowing measurements of supersaturated solution refractive indices to be made.

AOT measurements to determine supersaturated solution refractive indices were performed as follows. An aqueous solution of the organic solute of interest was nebulised in to the trapping cell and a single particle caught within the focussed laser beam. The RH within the trapping cell was then varied stepwise between 21% and 81% RH (citric acid), and 52 and 92% RH (malonic acid), with the particle allowed to equilibrate with its surroundings after each RH change until no further change in particle size was observed. Values of the particle radius, RI and wavelength dispersion in RI were estimated from Mie fits of the CERS spectra (centred at 650 nm) at each RH. All measured refractive indices were converted to their values at a wavelength of 589 nm using the wavelength dispersion retrieved in the fit and the method of Preston and Reid.[23-24] As the molar refraction mixing rule parameterises RI in terms of mass fraction of solute, the RH measured using the capacitance probe was converted in to a solute mass fraction using the group contribution model AIOMFAC. This allowed the AOT RI data to be compared directly with predictions from the molar refraction model.

Figures 11(a), (b) and (c) show the result of the AOT measurements and molar refraction mixing rule calculations for citric acid, tartaric acid and malonic acid respectively, with additional experimental data from Lienhard et al.[4] included for the first two compounds. Good agreement between the RI measurements, the data of Lienhard et al.[4] and molar refraction predictions at both sub-saturated and, importantly, supersaturated concentrations is observed for all three systems. For malonic acid, agreement between the experimental AOT data and the molar refraction RI predictions decreases at lower mass fractions of solute as this corresponds to measurements made at higher RH (>80%), where the accuracy of the capacitance RH probe is reduced. A particle containing aqueous malonic acid has a mass fraction of solute of ~ 0.55 at 80% RH.

In addition to measurements of the refractive indices of supersaturated solution droplets using the AOT, the pure component RIs for citric acid and tartaric acid were determined by nebulising an aqueous solution containing each organic and carbitol (2-(2-Ethoxyethoxy)ethanol) in a 1:3 molar ratio into the trapping cell at 80% RH. Once a particle was captured within the laser beam, the trapping cell RH was dropped immediately

to 0% RH. Carbitol is a volatile solvent and evaporates rapidly from the droplet, with its loss monitored through the band intensity of its characteristic spontaneous Raman peak at 628 nm. Citric acid is normally observed to effloresce under dry conditions preventing a determination of the value of the pure component melt RI by the AOT approach. However, we observed that the initial presence of carbitol in the droplet allowed the particle to equilibrate at 0% RH without crystallisation. The particle remained in the trap as a sub-cooled liquid droplet even after the complete evaporative loss of the characteristic Raman peak at 628 nm arising from carbitol. The RI of the remaining particle, containing only citric acid, could then be determined through CERS. As with citric acid, the addition of carbitol to a solution droplet containing tartaric acid allowed the pure tartaric acid particle to equilibrate at 0% RH while remaining as a liquid droplet. The pure component refractive indices of both compounds are compared with molar refraction mixing rule predictions in Figures 11(a) and 11(b), with good agreement between the experimental data and the model (citric acid predicted melt RI = 1.5038 , measured melt RI = 1.5086 ; tartaric acid predicted melt RI = 1.4966, measured melt RI = 1.5013).

6. Conclusion

It is of crucial importance in a range of applications and settings to represent accurately the physicochemical properties of aerosol particles of complex composition, both for understanding and predicting aerosol properties and processes, and for interpreting and comparing analysis methods. In this publication, we have considered in detail representations of the density and refractive index of organic aerosol. In particular, we have examined the dependence of these properties on RH and water content, considering the accuracy of mixing rules based on a comparison with sub-saturated solute concentration bulk data and with single particle measurements of droplets in a supersaturated solute state provided by the optical tweezers technique. We have concluded that the molar refraction mixing rule is most accurate for predictions of solution RI, with an accuracy that is determined by the accuracy of the treatment of solution density. If bulk density data is available to solute concentrations above a mass fraction of solute of 0.4, a third order polynomial parameterisation of solution densities is best able to represent the solution density data while allowing an accurate determination of the pure solute melt density. If the solubility limit of the organic component is less than a mass fraction of solute of 0.4, a density treatment that assumes ideal mixing is preferred over a polynomial fit. We have considered the properties of aqueous solutions of approximately 70 organic compounds including mono-, di- and tri-

carboxylic acids, alcohols, diols, nitriles, sulphoxides, amides, ethers, sugars, amino acids, aminium sulphates, and polyols. Although the uncertainties in the predictions of density and RI depend on the range of compositional data available for each compound, typical errors for solution density and refractive index are less than ±0.1 % and ±0.05 %, respectively. In future work, the use of the molar refraction mixing rule could be combined with group activity models for predicting the density[2] and refractive index of the pure organic components, with the latter based on estimates of molecular polarizability,[12] to provide a general framework for estimating the refractive indices of aqueous-organic aerosol of any composition and complexity. Of particular importance will be the extension of these approaches to predictions of the properties of aerosols of complex compositions containing the myriad of organic components found in ambient aerosol.

Supporting Information. Tables of measured refractive indices, densities and coefficients for polynomial fits to expressions described in the main text are included in the supporting information.

Acknowledgements

CC acknowledges support from the China Scholarship Council. REHM and JPR acknowledge support from the Natural Environment Research Council through grant NE/N006801/1. MIC acknowledges funding from NERC and the RSC through an Analytical Trust Fund studentship (NE/J01754X/1) and support from the Aerosol Society in the form of a CN Davies award. The EPSRC is acknowledged for studentship funding for AM and AMJR. GR acknowledges the Italian Ministry of Education for the award of a PhD studentship. YHZ acknowledges support from the NSFC through grant 91544223 and 21473009. Alex G. Cowling and Emily K. Hall are acknowledged for their contribution of density and refractive indices measurements used in this work.

References.

1 Clegg, S. L.; Wexler, A. S. Densities and Apparent Molar Volumes of Atmospherically Important Electrolyte Solutions. 1. The Solutes H_2SO_4, HNO_3, HCl, Na_2SO_4, $NaNO_3$, NaCl, $(NH_4)_2SO_4$, NH_4NO_3,

and NH_4Cl From 0 to 50 °C, Including Extrapolations to Very Low Temperature and to the Pure Liquid State, and $NaHSO_4$, NaOH, and NH_3 at 25 °C. *J. Phys. Chem. A* **2011**, *115*, 3393-3460.

2 Barley, M. H.; Topping, D. O.; McFiggans, G. Critical Assessment of Liquid Density Estimation Methods for Multifunctional Organic Compounds and their Use in Atmospheric Science. *J. Phys. Chem. A* **2013**, *117*, 3428-3441.

3 McMurry, P. H. A Review of Atmospheric Aerosol Measurements. *Atmos. Environ.* **2000**, *34*, 1959-1999.

4 Lienhard, D. M.; Bones, D. L.; Zuend, A.; Krieger, U. K.; Reid, J. P.; Peter, T. Measurements of Thermodynamic and Optical Properties of Selected Aqueous Organic and Organic ‐ Inorganic Mixtures of Atmospheric Relevance. *J. Phys. Chem. A* **2012**, *116*, 9954-9968.

5 Tang, I. N. Thermodynamic and Optical Properties of Mixed-Salt Aerosols of Atmospheric Importance. *J. Geophys. Res. Atmos.* **1997**, *102(2)*, 1883-1894.

6 Liu, Y.; Daum, P. H. Relationship of Refractive Index to Mass Density and Self-Consistency of Mixing Rules for Multicomponent Mixtures Like Ambient Aerosols. *J. Aerosol. Sci.* **2008**, *39*, 974-986.

7 Myhre, C.; Nielsen, C. J. Optical Properties in the UV and Visible Spectral Region of Organic Acids Relevant to Tropospheric Aerosols. *Atmos. Chem. Phys.* **2004**, *4*, 1759–1769.

8 Cai, C.; Stewart, D. J.; Reid, J. P.; Zhang, Y.; Ohm, P.; Dutcher, C. S.; Clegg, S. L. Organic Component Vapor Pressures and Hygroscopicities of Aqueous Aerosol Measured by Optical Tweezers. *J. Phys. Chem. A* **2015**, *119*, 704-718.

9 Girolami, G. S. A simple "Back of the envelope" method for estimating the densities and molecular volumes of liquids and solids. *J. Chem. Educ.* **1994,** *71*, 962.

10 Moise, T.; Michel Flores, J.; Rudich, Y. Optical Properties of Secondary Organic Aerosols and Their Changes by Chemical Processes. *Chem. Rev.* **2015**, *115*, 4400-4439.

11 Lambe, A. T.; Cappa, C. D.; Massoli, P.; Onasch, T. B.; Forestieri, S. D.; Martin, A. T.; Cummings, M. J.; Croasdale, D. R.; Brune, W. H.; Worsnop, D. R. et al. Relationship between Oxidation Level and Optical Properties of Secondary Organic Aerosol. *Environ. Sci. Technol.* **2013**, *47*, 6349-6357.

12 Cappa, C. D.; Che, D. L.; Kessler, S. H.; Kroll, J. H.; Wilson, K. R. Variations in Organic Aerosol Optical and Hygroscopic Properties upon Heterogeneous OH Oxidation. *J. Geophys. Res.* **2011**, *116*, D15204.

13 Zhang, X.; Massoli, P.; Quinn, P. K.; Bates, T. S.; Cappa, C. D. Hygroscopic Growth of Submicron and Supermicron Aerosols in the Marine Boundary Layer. *J. Geophys. Res. Atmos.* **2014**, *119*, 8384-8399.

14 Mason, B. J.; King, S.; Miles, R. E. H.; Manfred, K. M.; Rickards, A. M. J.; Kim, J.; Reid, J. P.; Orr-Ewing, A. J. Comparison of the Accuracy of Aerosol Refractive Index Measurements from Single Particle and Ensemble Techniques. *J. Phys. Chem. A* **2012**, *116*, 8547-8556.

15 Flores, J. M.; Zhao, D. F.; Segev, L.; Schlag, P.; Kiendler-Scharr, A.; Fuchs, H.; Watne, Å. K.; Bluvshtein1, N.; Mentel, Th. F.; Hallquist, M. et al. Evolution of the Complex Refractive Index in the UV Spectral Region in Ageing Secondary Organic Aerosol. *Atoms. Chem. Phys.* **2014**, *14*, 5793-5806.

16 Miles, R. E. H.; Rudic, S.; Orr-Ewing A. J.; Reid, J. P. The Influence of Uncertainties in the Diameter and Refractive Index of Calibration Polystyrene Beads on the Retrieval of Aerosol Optical Properties Using Cavity Ring Down Spectroscopy. *J. Phys. Chem. A* **2010,** *114*, 7077-7084.

17 Zarzana, K. J.; Cappa, C. D.; Tolbert, M. A. Sensitivity of Aerosol Refractive Index Retrievals Using Optical Spectroscopy. *Aerosol. Sci. Technol.* **2014**, *48*, 1133–1144.

18 Gharagheizi, F.; Ilani-Kashkouli, P.; Kamari, A.; Mohammadi, A. H.; Ramjugernath, D. Group Contribution Model for the Prediction of Refractive Indices of Organic Compounds. *J. Chem. Eng. Data.* **2014**, *59*, 1930-1943.

19 Redmond, H.; Thompson, J. E. Evaluation of a Quantitative Structure-Property Relationship (QSPR) for Predicting Mid-Visible Refractive Index of Secondary Organic Aerosol (SOA). *Phys. Chem. Chem. Phys.* **2011**, *13*, 6872-6882.

20 Dennis-Smither, B. J.; Marshall, F. H.; Miles, R. E. H.; Preston, T. C.; Reid, J. P. Volatility and Oxidative Aging of Aqueous Maleic Acid Aerosol Droplets and the Dependence on Relative Humidity. *J. Phys. Chem. A* **2014**, *118*, 5680-5691.

21 Krieger, U. K.; Marcolli, C.; Reid, J. P. Exploring the Complexity of Aerosol Particle Properties and Processes Using Single Particle Techniques. *Chem. Soc. Rev.* **2012**, *41*, 6631-6662.

22 Cai, C.; Stewart, D. J.; Preston, T. C.; Walker, J. S.; Zhang, Y. H.; Reid, J. P. A New Approach to Determine Vapour Pressures and Hygroscopicities of Aqueous Aerosols Containing Semi-Volatile Organic Compounds. *Phys. Chem. Chem. Phys.* **2014**, *16*, 3162-3172.

23 Preston, T. C.; Reid, J. P. Accurate and Efficient Determination of the Radius, Refractive Index, and Dispersion of Weakly Absorbing Spherical Particle Using Whispering Gallery Modes. *J. Opt. Soc. Am. B* **2013**, *30*, 2113-2122.

24 Preston, T. C.; Reid, J. P. Determining the Size and Refractive Index of Microspheres Using the Mode Assignments From Mie Resonances. *J. Opt. Soc. Am. A* **2015**, *32*, 2210.

25 Chu, K.; Thompson, A. R. Densities and Refractive Indices of Alcohol-Water Solutions n-Propyl, Isopropyl, and Methyl Alcohols. *J. Chem. Eng. Data.* **1962**, *3*, 358.

26 Acosta, J.; Arce, A.; Rodil, E.; Soto, A. A Thermodynamic Study On Binary and Ternary Mixtures of Acetonitrile, Water and Butyl Acetate. *Fluid. Phase. Equilibr.* **2002**, *203*, 83-98.

27 Granados, K.; Gracia-Fadrique, J.; Amigo, A.; Bravo, R. Refractive Index, Surface Tension, and Density of Aqueous Mixtures of Carboxylic Acids at 298.15 K. *J. Chem. Eng. Data.* **2006**, *51*, 1356-1360.

28 Lee, J. W.; Park, S. B.; Lee, H. Densities, Surface Tensions, and Refractive Indices of the Water + 1,3-Propanediol System. *J. Chem. Eng. Data.* **2000**, *45*, 166-168.

29 Al-Dujaili, A. H.; Yassen, A. A.; Awwad, A. M. Refractive Indices, Densities, and Excess Properties of 4-(2-Hydroxyethyl)morpholine + Water at (298.15, 308.15, and 318.15) K. *J. Chem. Eng. Data.* **2000**, *45*, 647-649.

30 Arce, A.; Blanco, A.; Soto, A.; Vidal, I. Densities, Refractive-Indexes, and Excess Molar Volumes of the Ternary-Systems Water Plus Methanol Plus 1-Octanol and Water Plus Ethanol Plus 1-Octanol and their Binary-Mixtures at 298.15 K. *J. Chem. Eng. Data.* **1993**, *38*, 336-340.

31 Lebel, R. G.; Goring, D. A. I. Density, Viscosity, Refractive Index, and Hygroscopicity of Mixtures of Water and Dimethyl Sulfoxide. *J. Chem. Eng. Data.* **1962**, *1*, 100.

32 Aminabhavi, T. M.; Gopalakrishna, B. Density, Viscosity, Refractive Index, and Speed of Sound in Aqueous Mixtures of NJV-Dimethylformamide, Dimethyl Sulfoxide, NJV-Dimethylacetamide, Acetonitrile, Ethylene Glycol, Diethylene Glycol, 1,4-Dioxane, Tetrahydrofuran, 2-Methoxyethanol, and 2-Ethoxyethanol at 298.15 K. *J. Chem. Eng. Data.* **1995**, 856.

33 Lide, D. R. *CRC Handbook of Chemistry and Physics*, 44th ed.; Chemical Rubber Company: Boca Raton, FL, 1962.

34 Topping, D.; Barley, M.; Bane, M. K.; Higham, N.; Aumont, B.; Dingle, N.; McFiggans, G. UManSysProp v1.0: an online and open-source facility for molecular property prediction and atmospheric aerosol calculations, *Geosci. Model Dev.* **2016,** *9*, 899-914.

35 Tang, I. N.; Munkelwitz, H. R. Water activities, densities, and refractive indices of aqueous sulfates and sodium nitrate droplets of atmospheric importance. *J. Geophys. Res.* **1994,** *99(D9)*, 18801–18808.

36 Tang, I. N. Chemical and size effects of hygroscopic aerosols on light scattering coefficients. *J. Geophys. Res.* **1996,** *1011(D14)*, 19245-19250.

37 Walker, J. S.; Wills, J. B.; Reid, J. P.; Wang, L.; Topping, D. O.; Butler, J. R.; Zhang, Y.-H. Direct comparison of the hygroscopic properties of ammonium sulfate and sodium chloride aerosol at relative humidities approaching saturation. *J. Phys. Chem. A* **2010,** *114*, 12682−12691.

38 Wills, J. B.; Butler, J. R.; Palmer, J.; Reid, J. P. Using optical landscapes to control, direct and isolate aerosol particles. *Phys. Chem. Chem. Phys.* **2009,** *11*, 8015−8020.

39 Mitchem, L.; Reid, J. P. Optical manipulation and characterisation of aerosol particles using a single-beam gradient force optical trap. *Chem. Soc. Rev.* **2008,** *37*, 756−769.

Figures

Figure 1. a) Percentage deviation and **b)** absolute deviation between measured refractive indices (n_{meas}) and predictions (n_{pred}) for binary aqueous mixtures of 21 different organics using the molar refraction mixing rule (black), the linear mass mixing rule (red) and the linear volume mixing rule (blue).

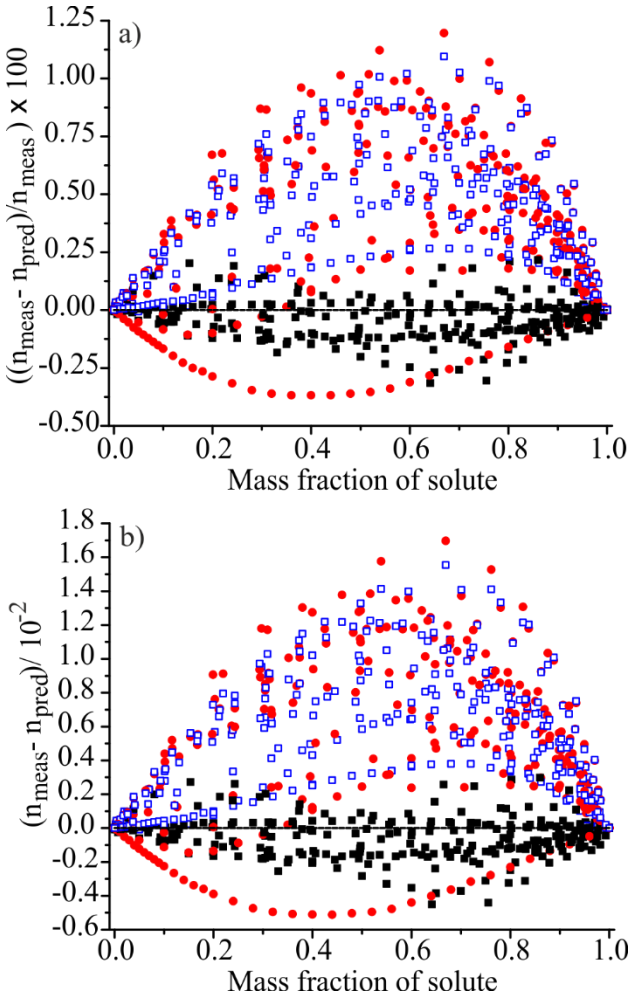

Figure 2. a) Percentage deviation and **b)** absolute deviation between measured solution densities (ρ_{meas}) and density predictions using the self-consistent ideal mixing rule (ρ_{IM}) for the 21 aqueous-organic mixtures. Each symbol represents a different system.

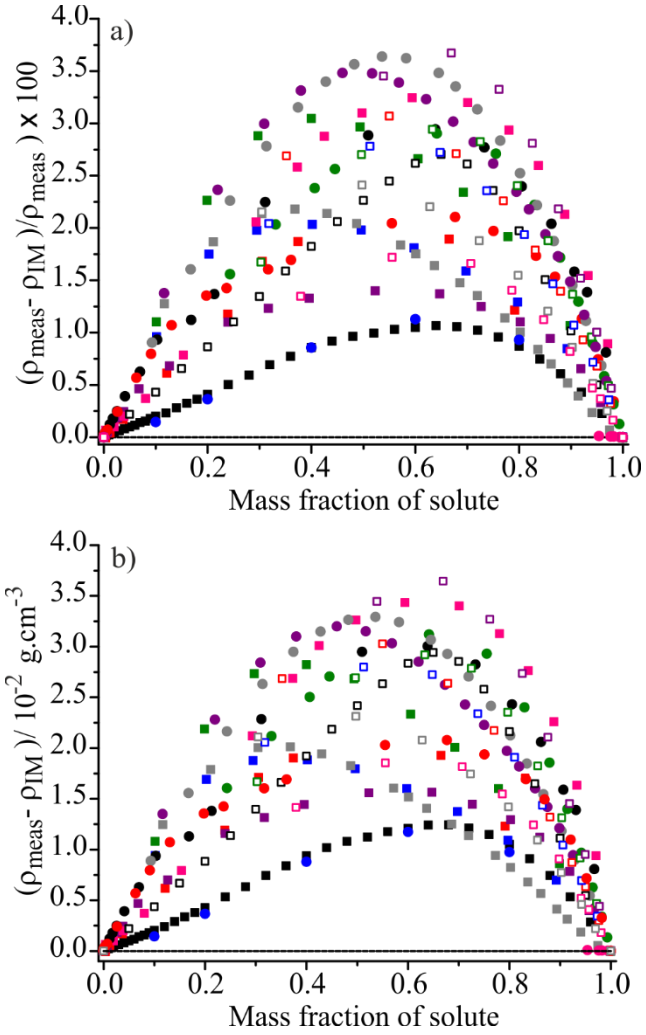

Figure 3. a) Percentage deviation and **b)** absolute deviation between the measured solution RI (n_{meas}) and predictions using the molar refraction mixing rule when using solution densities predicted by the self-consistent ideal mixing density rule (n_{IM-MR}). Each symbol represents a different system.

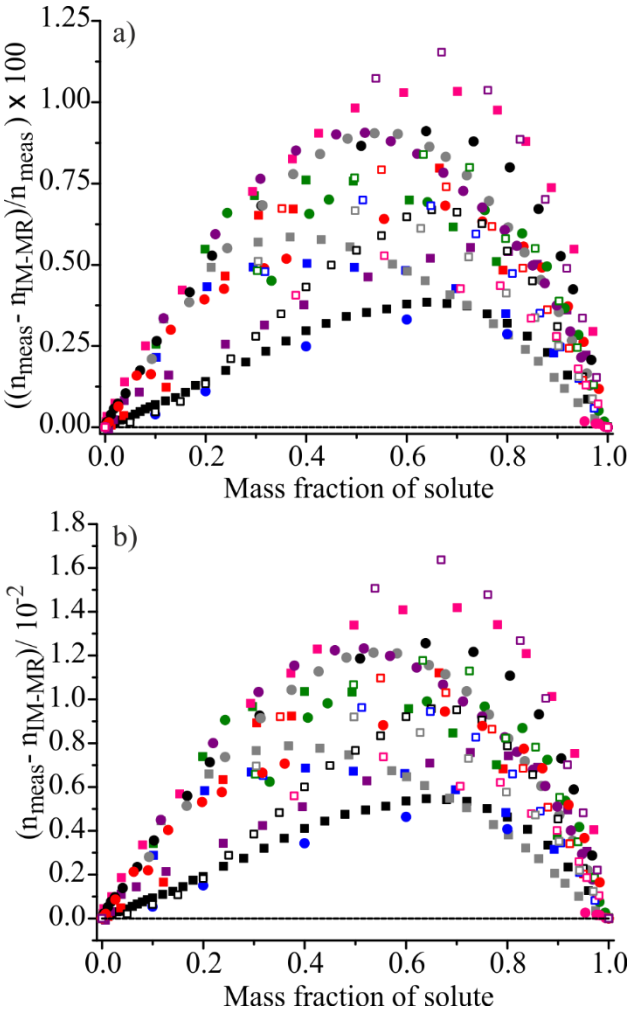

Figure 4. Correlation between the absolute error in the calculated mixture density using the ideal mixing rule and the resulting error in the RI predicted using the molar refraction mixing rule. The red line shows the result of a linear fit through the data, with the dashed lines representing the 95% confidence limits.

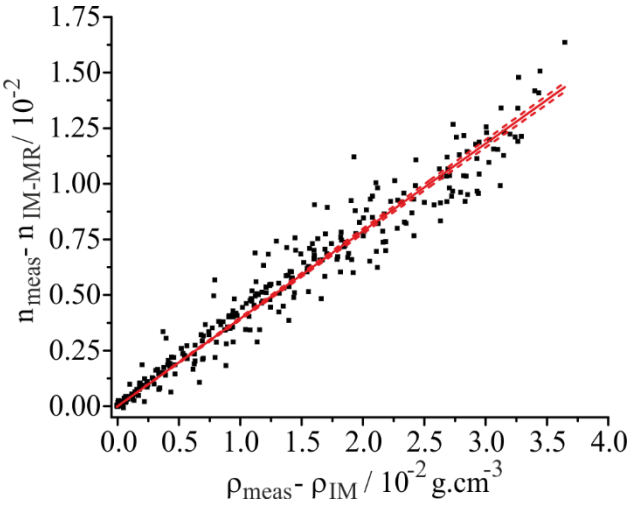

Figure 5. a) Fit of the self-consistent ideal mixing density rule to bulk data (black squares) for DL-Malic acid. The 95% confidence limits of the fit are shown by the width of the model line (red). **b)** Difference between experimentally measured density data (ρ_{meas}) and predictions from the self-consistent ideal mixing density rule (ρ_{IM}) for aqueous mixtures of 47 different organic compounds; black – amino acids; red – sugars; purple – aminium sulphates; green – mono/tricarboxylic acids, hydroxydicarboxylic acids, trichloroacetic acid, urea; blue – dicarboxylic acids. The black dashed lines correspond to the accuracy limit on measurements made using the density meter.

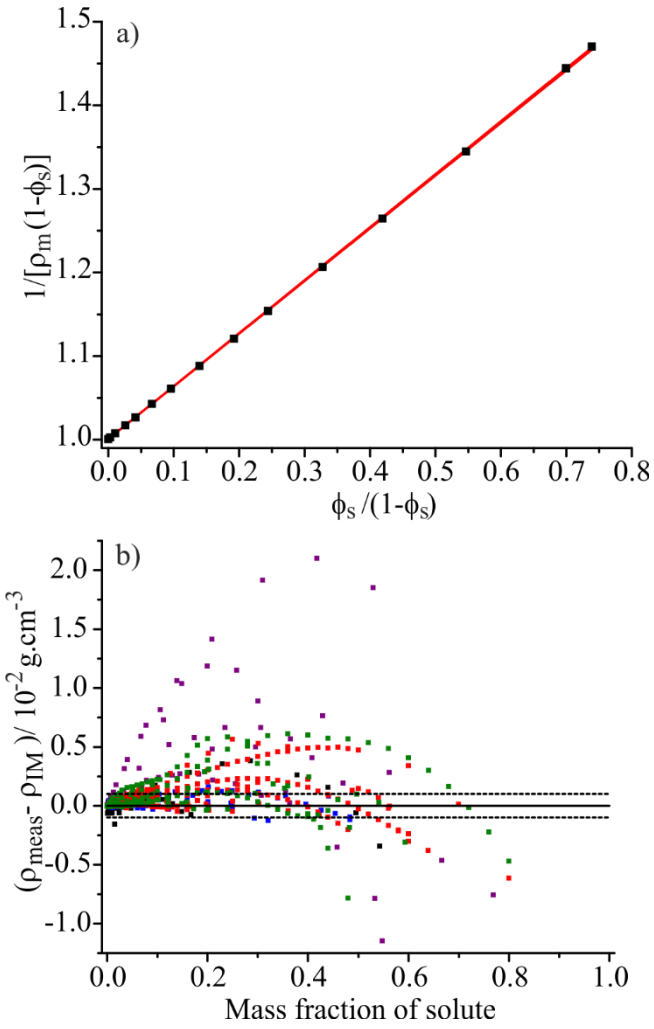

Figure 6. Apparent molar volume (V^{ϕ}) plotted as a function of the square root of the solute mass fraction ($\phi^{0.5}$) for a) urea, b) malonic acid and c) maltose.

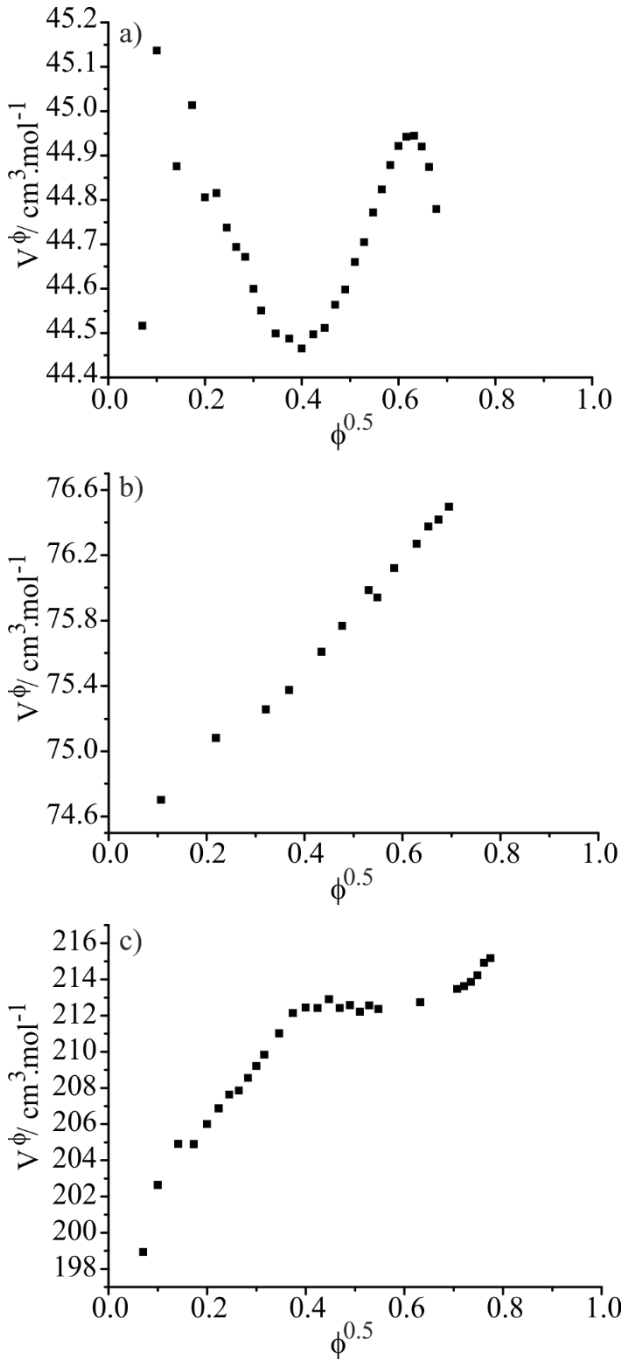

Figure 7. Difference between measured sub-saturated aqueous-organic mixture densities (ρ_{meas}) and predicted densities using **a)** second order, **b)** third order and **c)** fourth order polynomial fit to the data, as a function of the square root of solute mass fraction. Black – amino acids; red – sugars; purple – aminium sulphates; green – mono/tricarboxylic acids, hydroxydicarboxylic acids, tricholoroacetic acid, urea; blue – dicarboxylic acids. Dashed lines show the measurement uncertainty of the density meter.

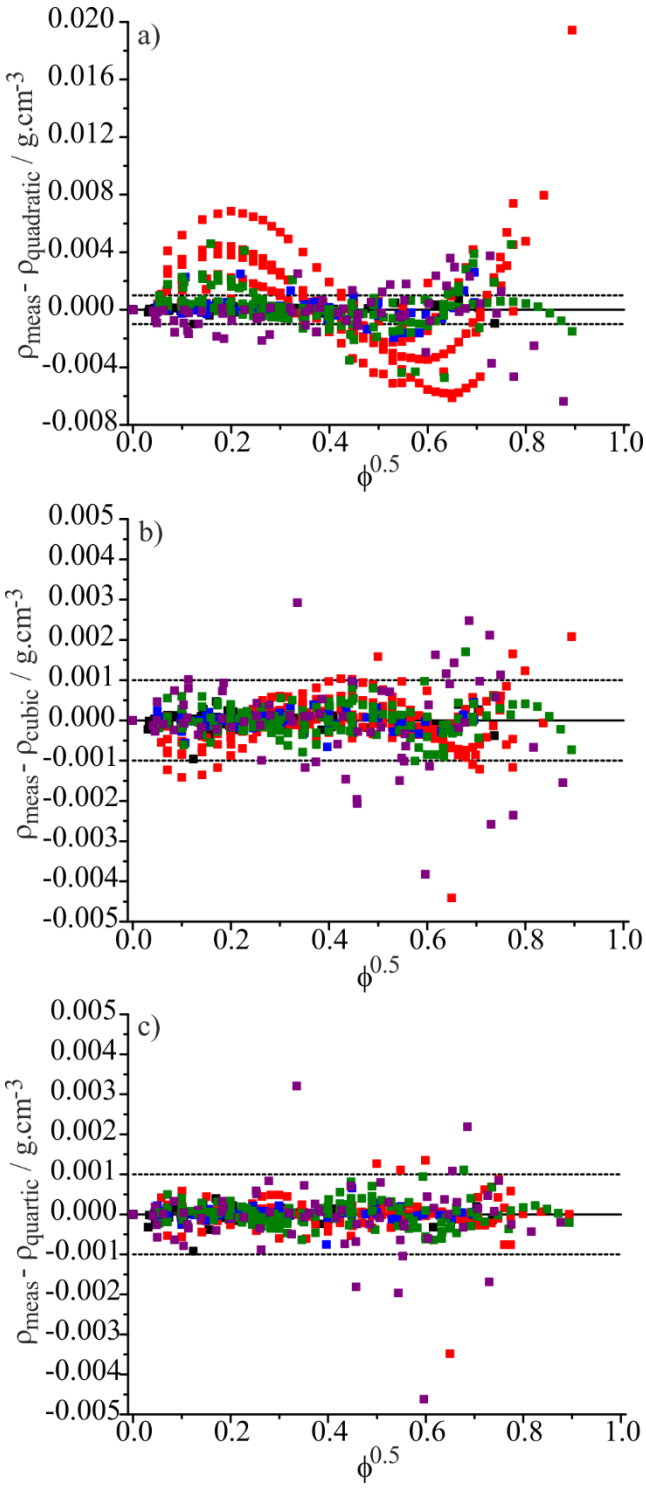

Figure 8. a) Uncertainty on extrapolated melt density (ρ_{melt}) estimated from different fitting procedures for the 47 compounds, plotted as a function of the maximum mass fraction of solute achievable in the bulk (blue – ideal mixing density treatment; black – third order polynomial fit; red – fourth order polynomial fit). **b)** Fractional agreement between the melt density predicted when all data is fitted and when a subset of data up to a given mass fraction of solute is fitted (Sucrose (black), Lactic acid (pink), MEAS (cyan), DEAS (orange), Sorbitol (red), Glucose (blue), Maltose (green), Citric acid (purple)). The dashed lines indicate a 2% error relative to the full fit value.

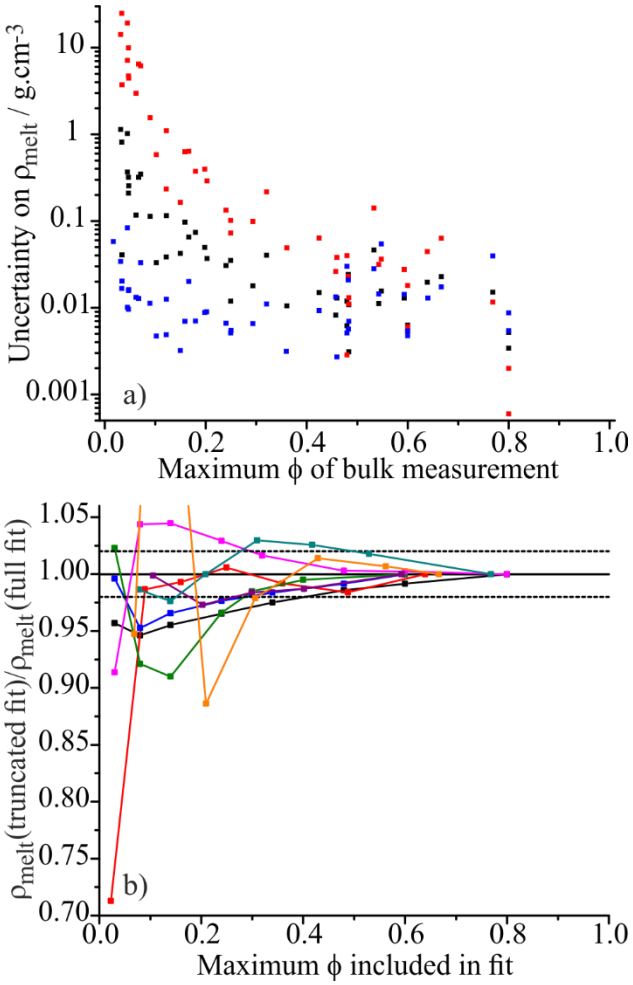

Figure 9. Comparison of solute melt densities predicted by a third order polynomial fit (black) and an ideal mixing density rule fit (red) to sub-saturated data, alongside predictions from UManSysProp (blue). Data are shown in terms of the maximum mass fraction of solute accessible in the bulk for **a)** dicarboxylic acids, **b)** sugars and **c)** mono/tricarboxylic and hydroxydicarboxylic acids.

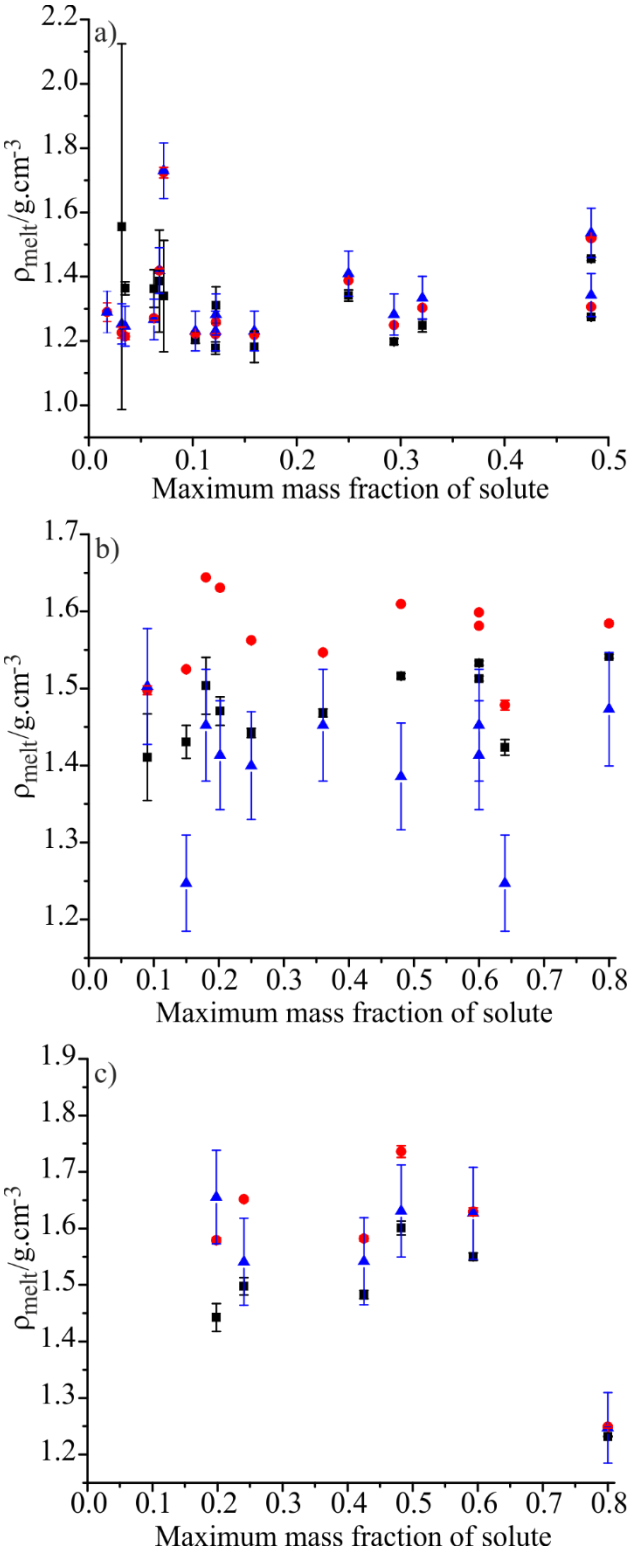

Figure 10. a) Difference between measured sub-saturated aqueous-organic mixture RIs (n_{meas}) and molar refraction mixing rule predicted RIs (n_{MR}), calculated using the best fit density treatment for each of the 47 compounds (black – amino acids; red –sugars; purple – aminium sulphates; green – mono/tricarboxylic acids, hydroxydicarboxylic acids, trichloroacetic acid, urea; blue – dicarboxylic acids). **b)** Uncertainty in the melt RI predicted using the molar refraction mixing rule, plotted as a function of the maximum mass fraction of solute achievable in the bulk.

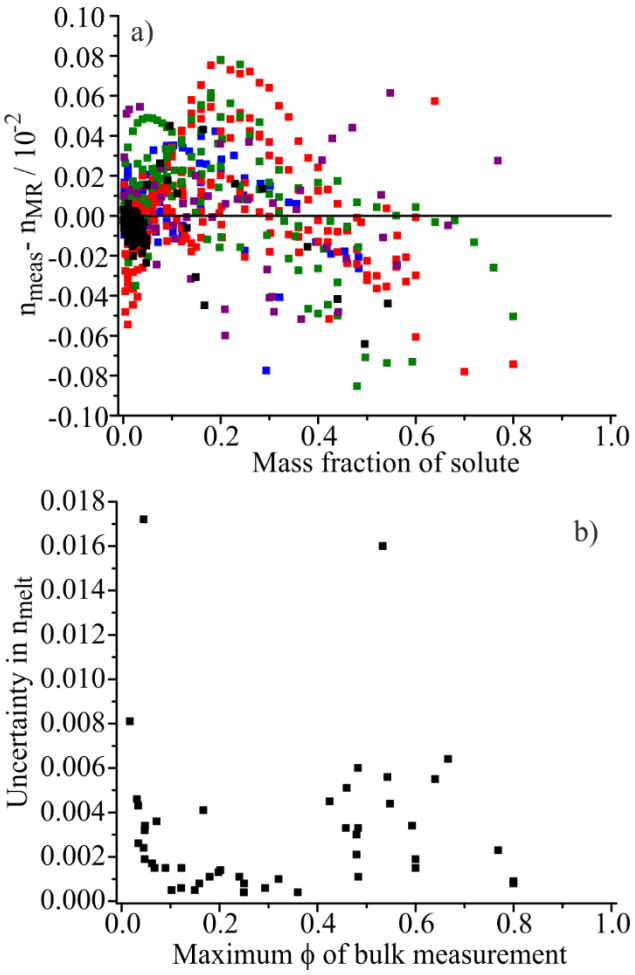

Figure 11. Comparison of measured supersaturated solution refractive indices for aqueous mixtures of **a)** citric acid, **b)** tartaric acid and **c)** malonic acid with molar refraction mixing rule RI predictions (black line), based on a fit to sub-saturated RI data (CRC handbook – black circles; this study – empty squares). The red lines show the 95% confidence level in the molar refraction predicted RI and arises from the uncertainty in the melt density. AOT measured supersaturated RIs are shown by the black squares; literature data from Lienhard et al. are shown as empty circles. The red circles show the measured pure component melt RI in the AOT.

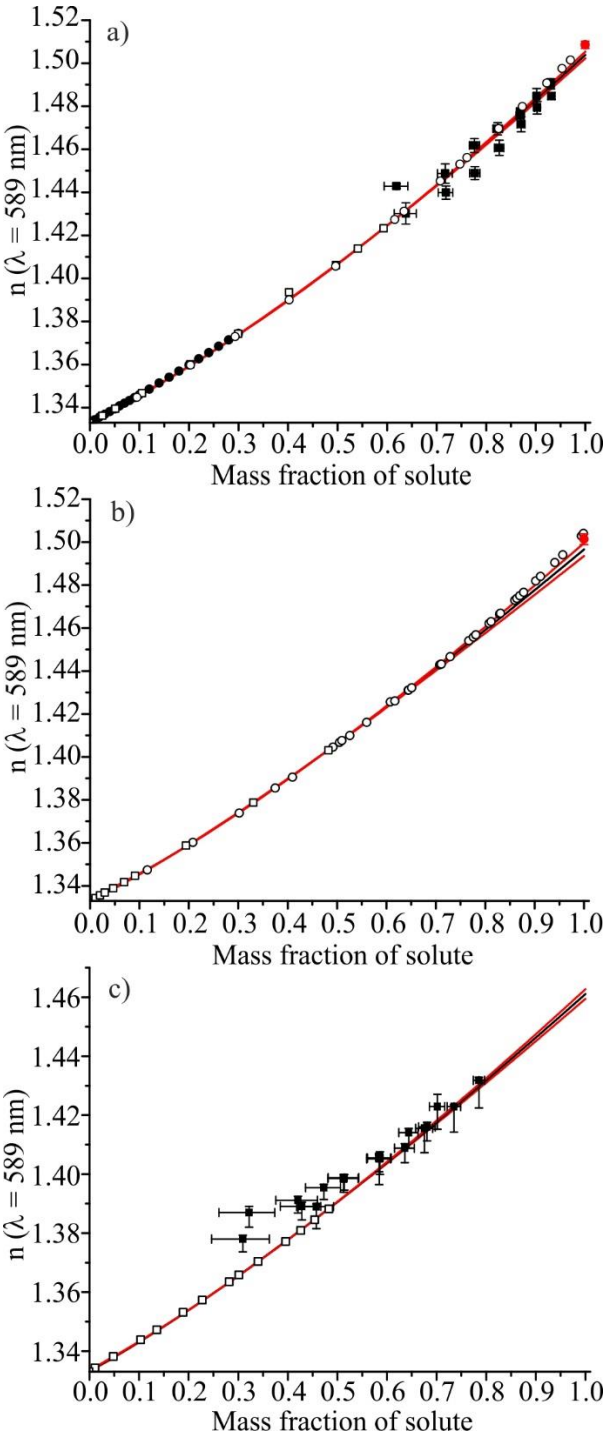

Table of Contents Graphic